

Artificial Wrestling: A Dynamical Formulation of Autonomous Agents Fighting in a Coupled Inverted Pendula Framework

Katsutoshi YOSHIDA*, Shigeki MATSUMOTO* and Yoichi MATSUE**

* Department of Mechanical and Intelligent Engineering, Utsunomiya University

7-1-2 Yoto, Utsunomiya, Tochigi, 321-8585, Japan

E-mail: yoshidak@cc.utsunomiya-u.ac.jp

** Yuhara Mfg Co., Ltd

1256 Ujiie, Sakura, Tochigi 329-1311, Japan

Abstract

We develop autonomous agents fighting with each other, inspired by human wrestling. For this purpose, we propose a coupled inverted pendula (CIP) framework in which: 1) tips of two inverted pendulums are linked by a connection rod, 2) each pendulum is primarily stabilized by a PD-controller, 3) and is additionally equipped with an intelligent controller. Based on this framework, we dynamically formulate an intelligent controller designed to store dynamical correspondence from initial states to final states of the CIP model, to receive state vectors of the model, and to output impulsive control forces to produce desired final states of the model. Developing a quantized and reduced order design of this controller, we have a practical control procedure based on an off-line learning method. We then conduct numerical simulations to investigate individual performance of the intelligent controller, showing that the performance can be improved by adding a delay element into the intelligent controller. The result shows that the performance depends not only on quantization resolutions of learning data but also on delay time of the delay element. Finally, we install the intelligent controllers into both pendulums in the proposed framework to demonstrate autonomous competitive behavior between inverted pendulums.

Key words : Multiagent System, Competitive Problem, Intelligent Control, Nonlinear Dynamics, Reachable Set

1. Introduction

Wrestling seems to be composed artificially of two mechanical agents maintaining their balance, coupled via mechanical interactions such as contact, connection, collision, etc., and equipped with intelligent controllers competitive with each other. In this paper, we develop a simple model to create such competitive agents. For this purpose, we propose a coupled inverted pendula (CIP) framework in which: 1) tips of two inverted pendulums are linked by a connection rod, 2) each pendulum is primarily stabilized by a PD-controller, 3) and is additionally equipped with an intelligent controller that individually generates a series of impulsive internal forces to achieve its own desired final states based on knowledge of correspondence from initial states to the final states.

In general, multiple agents can exhibit competitive and cooperative dynamics when sharing common resources and environments. Historically, early mathematical insights into such mutual interactions seem to have appeared in the field of mathematical ecology (Hofbauer and Sigmund, 1998) in which population dynamics of different species sharing a common environment is described by a system of coupled nonlinear differential equations such as the Lotka-Volterra equation. Contrary to the ecosystem in which the medium of interaction is given by environments, in our CIP framework, the medium of interaction is given by a mechanical structure. In our previous study (Yoshida and Ohta, 2008), we have already demonstrated that the CIP model, even without intelligent controllers, can produce competitive dynamics comparable to that in the ecosystem such as coexistence and dominance by assigning competitive meanings to the stable equilibriums of the CIP model. Although quite similar mechanical models have been considered in the field of multiple manipulator systems (Hsu, 1993; Nakamura et al., 1987; Panwar et al., 2012), they have only focused on cooperative dynamics because of their aim at developing coordinated motions in those systems.

* Received DD MM, YYYY (No. XX-XXXX)

[DOI: 10.1299/XXX.X.1]

Copyright © 2014 by JSME

In our study mentioned above (Yoshida and Ohta, 2008), each pendulum is PD-controlled to be bistable at the top and bottom dead points such that the coupled system produces quadra-stability. In absence of additional inputs, this system converges into one of the four stable positions (equilibriums) depending upon the initial conditions. We then applied a single impulsive force to one of the pendulums to generate switching behavior from a given stable position to a desired stable position and considered it as a prototype of fighting-like behavior. In this prototype, however, the behavior is exactly determined by the initial position and strength of the impulse because of uniqueness of solution of differential equation. Therefore, it is quite hard to say that this first prototype is comparable to the wrestling players who seek how to generate internal forces to achieve desired final positions in autonomous ways.

In order to build such autonomous agents fighting with each other, a certain intelligent motion controller is required. On such controllers, extensive research has been conducted in the field of multi-robot systems (Maravall et al., 2013; Pagello et al., 1999). The major issue in this field appears to be how to obtain cooperative group dynamics of robots both in algorithm-based approaches (Stone and Veloso, 2000) and in differential-equation-based approaches (Hsu, 1993; Nakamura et al., 1987; Panwar et al., 2012).

On the other hand, competitive group behavior seems to have been studied mainly based on algorithm-based approaches. For example, Nelson et al. (2004) studied an evolutionary controller to investigate a form of reinforcement learning that makes use of competitive tournaments of games (robot capture the flag) played by individuals in a population of neural controllers. Moreover, Wu et al. (2013) developed rule or knowledge-based techniques to analyze strategy in robot soccer game. In this way, in contrast to our approach (Yoshida and Ohta, 2008), differential-equation-based techniques are not always essential in these studies because they seek step-by-step algorithms predicting free space determined by other robots.

In this paper, we first introduce the CIP framework to describe the competitive behavior in differential-equation-based manners. Next, we develop a competitive intelligent controller that receives state vectors of the CIP model and outputs impulsive control forces to produce desired final states. After evaluating individual performance of the intelligent controller and investigating how to improve the performance, we will demonstrate autonomous competitive behavior between two inverted pendulums equipped with the proposed intelligent controllers.

2. Coupled inverted pendula framework

2.1. Coupled inverted pendula

In order to create wrestler-like mechanical agents maintaining their balance while being coupled mechanically with each other, we consider a CIP model (Yoshida and Ohta, 2008) as shown in Fig. 1. Each inverted pendulum consists of a cart moving along the horizontal floor ($Y = 0$) and a simple pendulum rotating about a point on the cart. For simplicity, a common physical specification is given to both of the pendulums where m_θ is a mass and r is a length of the pendulum, and m_x is a mass of the cart. Linking the tips of the pendulums with a viscoelastic connection rod of length w , we obtain the CIP model in Fig. 1 where T_i is an input torque on θ_i , f_i is a reaction force acting on the tip of i th pendulum, and k_w and c_w are a spring coefficient and a viscous friction coefficient of the rod respectively. We assume that a mass of the rod is negligible. As configuration of this linkage is uniquely determined by the four variables: horizontal displacements of the carts x_1, x_2 and slant angles of the pendulums θ_1, θ_2 , the dynamics of this linkage is described by the eight-dimensional state vector:

$$\mathbf{x} := (\mathbf{x}_1^T, \mathbf{x}_2^T)^T, \quad \mathbf{x}_i := (x_i, \dot{x}_i, \theta_i, \dot{\theta}_i)^T \quad (i = 1, 2), \quad (1)$$

where A^T denotes the transpose of a matrix A . According to Lagrangian mechanics and assuming viscous friction forces $c_x \dot{x}_i$ and $c_\theta \dot{\theta}_i$ on x_i and θ_i respectively, we obtain equations of motion (EOM) of the CIP model in Fig. 1 as follows:

$$\begin{cases} (m_x + m_\theta) \ddot{x}_i + (m_\theta r \cos \theta_i) \ddot{\theta}_i - m_\theta r \dot{\theta}_i^2 \sin \theta_i = -c_x \dot{x}_i + (1, 0) f_i, \\ (m_\theta r \cos \theta_i) \ddot{x}_i + (m_\theta r^2) \ddot{\theta}_i - m_\theta g r \sin \theta_i = -c_\theta \dot{\theta}_i + r(\cos \theta_i, -\sin \theta_i) f_i + T_i \end{cases} \quad (i = 1, 2), \quad (2)$$

where $\dot{X} := dX/dt$.

2.2. Reaction force from the connection rod

We calculate the reaction force, say \mathbf{p} , from the connection rod as shown in Fig. 2. The displacement vector \mathbf{w} from

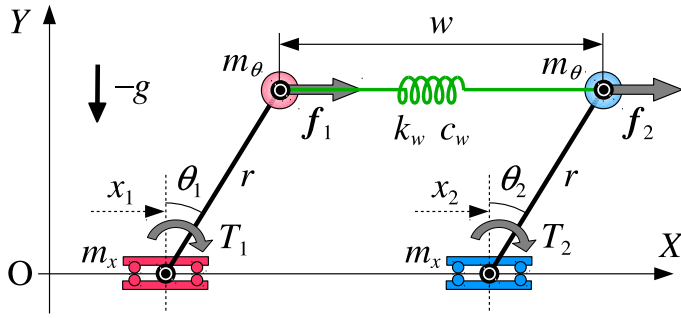
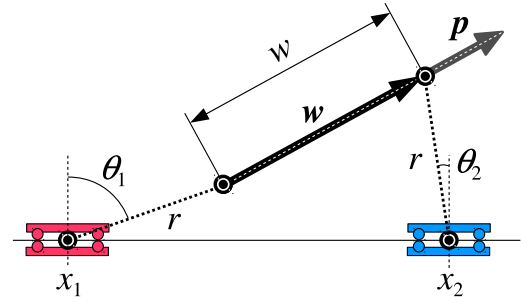


Fig. 1 Coupled inverted pendula model via viscoelastic connection.

Fig. 2 Reaction force \mathbf{p} from the connection rod.

the left-hand tip to the right-hand tip of pendulums is given by

$$\mathbf{w} = \begin{bmatrix} w_X \\ w_Y \end{bmatrix} := \mathbf{X}_2 - \mathbf{X}_1, \quad \mathbf{X}_i = \begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} x_i + r \sin \theta_i \\ r \cos \theta_i \end{bmatrix} \quad (i = 1, 2), \quad (3)$$

and the length of rod is expressed as

$$w = \|\mathbf{w}\| = \sqrt{\{(x_2 - x_1) + r(\sin \theta_2 - \sin \theta_1)\}^2 + \{r(\cos \theta_2 - \cos \theta_1)\}^2} \quad (4)$$

with the time derivative $\dot{w} = (\dot{w}_X w_X + \dot{w}_Y w_Y)/w$. Then, we model viscoelasticity of the connection rod as

$$\mathbf{p} = \|\mathbf{p}\| := -k_w(w - w_0) - c_w \dot{w}, \quad (5)$$

where w_0 is a natural length of the connection rod. As the force vector \mathbf{p} is parallel to the displacement vector \mathbf{w} , we have the reaction force:

$$\mathbf{p} = (p/w)\mathbf{w}. \quad (6)$$

Substituting \mathbf{p} into the EOM (2) through,

$$\mathbf{f}_1 = -\mathbf{p}, \quad \mathbf{f}_2 = \mathbf{p} \quad (\text{or } \mathbf{f}_i = (-1)^i \mathbf{p}), \quad (7)$$

we obtain an analytic expression of the CIP model shown in Fig. 1 via the viscoelastic connection.

2.3. Modeling floor

In the previous study (Yoshida and Ohta, 2008), the pendulum can fall down freely to the bottom dead point, in other word, there was no floor in the previous CIP model. In that case, both of forward and backward falling motions converge to the same equilibrium of the model so that orbital information is required in order to detect the direction of falling. Not only to simplify the detection process but also to develop more realistic simulator of wrestling, we introduce the floor model to the CIP model in the following manner.

Based on penalty methods (Moore and Wilhelms, 1988), we first model a normal force R_i from the floor ($Y = 0$) acting on the tip of i th pendulum as

$$R_i = U(-Y_i)\{-k_f Y_i - c_f \dot{Y}_i\}, \quad (8)$$

where Y_i is a height of the i th tip from the floor in (3), $U(\cdot)$ is a unit step function, and k_f, c_f are viscoelastic parameters representing property of the reaction. In practice, in order to avoid numerical errors, we approximate the step function with a sigmoid function differentiable, defined by

$$U_\sigma(s) := \{1 + \exp(-\sigma s)\}^{-1}, \quad (9)$$

where $\lim_{\sigma \rightarrow \infty} U_\sigma(s) = U(s)$ holds.

Furthermore, a Coulomb friction force F_i acting on the i th tip from the floor can be expressed as

$$F_i = -\mu R_i \operatorname{sgn}(\dot{X}_i), \quad (10)$$

where μ is a friction coefficient, \dot{X}_i is a relative horizontal velocity of the i th tip from the floor, and $\operatorname{sgn}(\cdot)$ is a unit signum function whose smooth approximation can be given by $\operatorname{sgn}(s) \approx \operatorname{sgn}_\sigma(s) := 2U_\sigma(s) - 1$.

Therefore, the CIP model via the viscoelastic connection in (7) on the floor can be obtained by substituting

$$\mathbf{f}_i = (-1)^i \mathbf{p} + (F_i, R_i)^T \quad (i = 1, 2) \quad (11)$$

into the EOM in (2).

| $\theta_1 \backslash \theta_2$ | 0 | $-\pi/2$ | $\pi/2$ |
|--------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
| 0 | ω_1 : draw | ω_2 : left win pulling | ω_3 : left win pushing |
| $\pi/2$ | ω_4 : right win pulling | ω_5 : draw | ω_6 : draw |
| $-\pi/2$ | ω_7 : right win pushing | ω_8 : draw | ω_9 : draw |

Fig. 3 Competitive interpretation of the equilibriums.

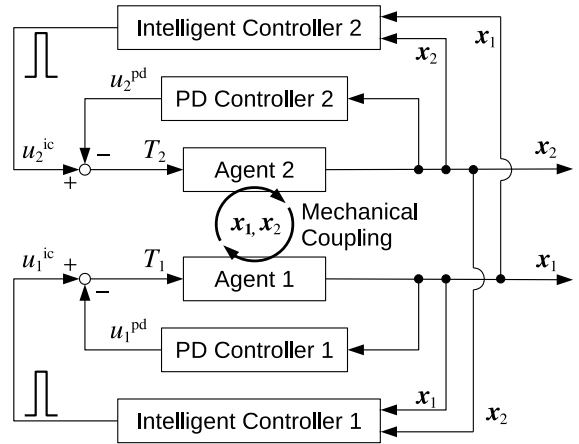


Fig. 4 Architecture of the proposed control system to generate competitive motions.

2.4. Standing control with falling

The CIP framework in absence of intelligent controllers is completed by giving dynamical meanings of winning and losing to states of the CIP model. To this end, we begin with developing a feedback controller by which each inverted pendulum on the floor forms three stable equilibriums: $\theta_i = 0$ for standing or winning and $\theta_i = \pm\pi/2$ for falling or losing. This can be done by introducing a feedback controller in the following form:

$$T_i = u_i^{\text{pd}} := \text{trap}_\alpha(\theta_i; \Delta\theta) \{-K_p\theta_i - K_d\dot{\theta}_i\} \quad (i = 1, 2), \quad (12)$$

where

$$\text{trap}_\alpha(\theta; \Delta\theta) := U_\alpha(\theta + \Delta\theta) \cdot U_\alpha(-\theta + \Delta\theta) \quad (13)$$

is a smooth trapezoidal function of unit height and centered at $\theta = 0$ made of a product of the sigmoid function in (9), $\Delta\theta > 0$ is a half width of the trapezoidal shape, and the shape is getting steeper as α increases.

It follows from the deadband characteristics in (12) that u_i^{pd} simply acts as a PD controller within the limit $|\theta_i| < \Delta\theta$ while it rapidly cuts off the output outside of the interval. Therefore, appropriate setting of the gains K_p, K_d make it possible for the i th pendulum to be stabilized about the standing position $\theta_i = 0$ while to be falling down to the floor when $|\theta_i|$ exceeds the given limit $\Delta\theta$.

It is worthy to note that in the field of gerontology and related fields, human standing (or falling) limits comparable to the threshold $\Delta\theta$ have been measured by the *functional reach test* (Duncan et al., 1990) in which the difference between arm length and maximal forward reach of human subjects is measured to evaluate risk of falls of them.

2.5. CIP framework

In what follows, we compare the inverted pendulums in this model to wrestler-like agents maintaining their standing balance. Since each agent with the standing control in (12) has the three stable equilibriums, a pair of the agents being coupled with each other under the suitable conditions can produce $3 \times 3 = 9$ stable equilibriums:

$$\omega_i := \lim_{t \rightarrow \infty} \mathbf{x}(t) = (\bar{x}_1, 0, \bar{\theta}_1, 0, \bar{x}_2, 0, \bar{\theta}_2, 0)^T \quad (i = 1, \dots, 9), \quad (14)$$

as shown in Fig. 3 when equating horizontal translations of final position \bar{x}_1, \bar{x}_2 without loss of generality. Namely, the components $x_1(t), x_2(t)$ of the solution $\mathbf{x}(t)$ of (2) are not stable asymptotically but neutrally because no restoring forces on $x_1(t), x_2(t)$ are assumed by definition. It also should be noted that due to the penalty method in (8), the gravity force makes the equilibrium $\bar{\theta}_i$ on the floor slightly exceed the floor, i.e., $|\bar{\theta}_i| - \pi/2 > 0$, but we formally denote $\theta_i = \bar{\theta}_i = \pm\pi/2$ because this slight exceedance only affects almost converged states and does not change the correspondence from the initial to final states.

We then attach competitive meanings to the nine equilibriums as listed in Fig. 3 in which the agent that remains standing is regarded as a winner. Eventually, we have the *CIP framework* composed of the set of (A) and (B):

- (A) The CIP model: the system of equations defined in (2), (11) and (12).
- (B) The win-loss matrix: the competitive interpretation of the nine equilibriums defined in Fig. 3.

3. Intelligent controller

Based on the CIP framework in Section 2.5, we develop an intelligent controller (IC) to produce desired final positions in Fig. 3 from given initial states of CIP by generating certain impulsive forces.

3.1. Problem setting and requirements

According to the definition of the state vector $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T$ in (1), the CIP model in (2) can be expressed as an eight-dimensional dynamical system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{T}), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{T} := (T_1, T_2)^T \quad (15)$$

that can be divided into a pair of four-dimensional subsystems:

$$\begin{cases} \dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, T_1), \\ \dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_2, \mathbf{x}_1, T_2). \end{cases} \quad (16)$$

We introduce the IC by adding an intelligent control input u_i^{ic} to the torque T_i of CIP model as

$$\mathbf{T} := \mathbf{u}^{\text{pd}} + \mathbf{u}^{\text{ic}} = (u_1^{\text{pd}}, u_2^{\text{pd}})^T + (u_1^{\text{ic}}, u_2^{\text{ic}})^T, \quad (17)$$

where u_i^{pd} is the standing control input already given in (12).

In the present study, the input-output relationship around u_i^{ic} is designed as shown in Fig. 4, in which u_i^{ic} receives all the state vectors $\mathbf{x}_1, \mathbf{x}_2$ and outputs a series of impulsive forces given by

$$u_i^{\text{ic}}(t) := \sum_{j=1}^N P_i I_{\Delta\tau}(t - t_i^j), \quad (18)$$

where

$$I_{\Delta\tau}(t) = \begin{cases} (\Delta\tau)^{-1} & (0 \leq t < \Delta\tau), \\ 0 & (\text{otherwise}) \end{cases} \quad (19)$$

is a rectangular function of unit area of width $\Delta\tau \ll 1$, P_i is an angular impulse of the input torque $u_i^{\text{ic}}(t)$, and $\{t_i^1, \dots, t_i^N\}$ is a series of rise time satisfying

$$t_i^1 < t_i^2 < \dots < t_i^N, \quad \max_{j,k} |t_i^j - t_i^k| \geq \tau_G \geq \Delta\tau, \quad (20)$$

where τ_G is a relaxation time to avoid overlapped outputs.

In practical implementation, the rise times t_i^1, \dots, t_i^N are supposed to be determined sequentially by a real time architecture described in Fig. 5, which is composed of three components, a classifier C , a selector $S_{\mathbb{J}}$, and an impulse generator G .

3.2. Classifier C

We define the classifier C as a function from a state vector $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T$ at the time $t = t_0$, say $\mathbf{x}(t_0) = \xi_0$, to an index number ν of equilibrium ω_ν . The function C takes the value $C(\xi_0) = \nu$ if a solution of the following system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{T}), \quad \mathbf{x}(t_0) = \xi_0, \quad \mathbf{T} := \mathbf{u}^{\text{pd}} + \begin{bmatrix} P_1 I_{\Delta\tau}(t - t_0) \\ 0 \end{bmatrix} \quad (21)$$

converges to the equilibrium ω_ν . From this definition in which a single impulse at $t = t_0$ is applied only on the left-hand agent, it is implied that the classifier C is valid only in absence of additional inputs, in other words, it can fail to return correct equilibriums for more general cases of input as in (18) where both agents can produce impulsive forces for their own decisions. Despite that, we will proceed with a discussion to build a first prototype of artificial wrestling.

Consider the transition operator of a solution of (21) as

$$\mathbf{x}(t) := \phi_t(\xi_0, \mathbf{T}), \quad (22)$$

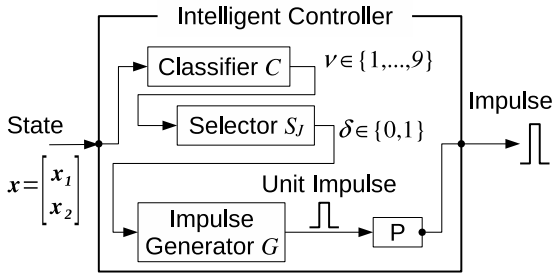


Fig. 5 Intelligent controller.

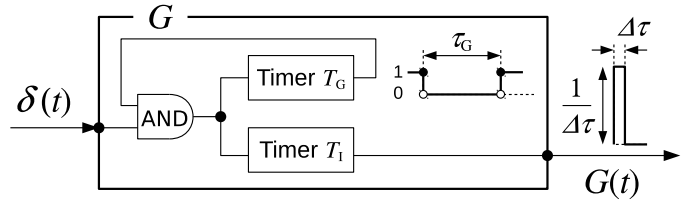


Fig. 6 Impulse generator.

and define a set of the initial state ξ_0 approaching the equilibrium ω_i as

$$\Phi_i := \left\{ \xi_0 \in R^8 \mid \lim_{t \rightarrow \infty} \phi_t(\xi_0, T) = \omega_i \right\}. \quad (23)$$

The set Φ_i is generally called a basin of attraction (Lhommeau et al., 2011) or a reachable set (Bayadi et al., 2013). It follows from uniqueness of solution of initial value problem in (21) that the reachable set in (23) satisfies,

$$\Phi_i \cap \Phi_j = \emptyset \quad (i \neq j). \quad (24)$$

Thus, the classifier C is obtained as a single-valued function:

$$C(\xi_0) := \nu \quad \text{if} \quad \xi_0 \in \Phi_\nu. \quad (25)$$

In Section 3.5, we will discuss a numerical approximation of C because explicit expressions of $C(\cdot)$ are hardly obtainable from nonlinear systems such as (15) and also (21).

3.3. Selector $S_{\mathbb{J}}$

In the competitive problem in Fig. 3, some of the equilibriums are selected depending upon strategies of the agent considered. Such a selection process can be modeled by a selector $S_{\mathbb{J}}$ given by

$$\delta = S_{\mathbb{J}}(\nu) := \begin{cases} 1 & (\nu \in \mathbb{J} \subset \{1, \dots, 9\}), \\ 0 & (\text{otherwise}), \end{cases} \quad (26)$$

where $\nu = C(x)$ is an output of the classifier and \mathbb{J} is a given subset of indices of the equilibriums $\omega_1, \dots, \omega_9$. For example, the trajectory $x(t)$ in (21) starting from an initial state ξ_0 satisfying

$$(S_{\mathbb{J}} \circ C)(\xi_0) := S_{\mathbb{J}}(C(\xi_0)) = 1 \quad \text{for} \quad \mathbb{J} = \{2, 3\} \quad (27)$$

converges one of the two equilibriums ω_2 and ω_3 .

3.4. Impulse generator G

The impulse generator G is designed to receive the binary signal $\delta(t) = S_{\mathbb{J}}(\nu(t))$ from the selector and output the unit impulse $G(t)$ as shown in Fig. 6, which is composed of a two-input AND gate and two timer functions T_l and T_G . The timer T_l produces a unit impulse as

$$G(t) = T_l(t) := I_{\Delta\tau}(t - t_r), \quad (28)$$

and the timer T_G cuts off the binary signal $\delta(t) = S_{\mathbb{J}}(\nu(t))$ for a given relaxation time τ_G discussed in (20) as

$$T_G(t) := \begin{cases} 0 & (t_r < t < t_r + \tau_G), \\ 1 & (\text{otherwise}), \end{cases} \quad (29)$$

where t_r is the rise time from 0 to 1 of the Boolean product $\hat{\delta}(t) = S_{\mathbb{J}}(\nu(t)) \wedge T_G(t)$. Note that although the signal $\hat{\delta}(t)$ can be pulses of infinitesimal width in this continuous time expression, this problem will never arise in discrete time applications with digital computers.

3.5. Numerical approximation of the classifier C

Explicit expression of $C(\mathbf{x})$ is hardly obtainable from nonlinear systems such as (15). It seems that there are two types of solutions: solving (15) in process or making a numerical table of the mapping $C : \mathbf{x} \mapsto \nu$ in advance. We take the latter approach in the following manner.

For practical applications, we introduce a linear measurement equation:

$$\mathbf{y} = H\mathbf{x}, \quad \mathbf{y} \in R^M, \mathbf{x} \in R^8, \quad (30)$$

where $M \leq 8$ and H is a $M \times 8$ matrix of rank M , and define a reduced-order reachable set in the following sense:

$$\Psi_i = H(\Phi_i) := \left\{ \eta_0 \in R^M \mid \xi_0 = h(\eta_0) \in R^8, \lim_{t \rightarrow \infty} \phi_t(\xi_0, T) = \omega_i \right\}, \quad (31)$$

where $\mathbf{x}' = h(\mathbf{y})$ ($\neq \mathbf{x}$ in general) is a certain inverse of $\mathbf{y} = H\mathbf{x}$. An actual example of h will be given in Section 3.6.

Firstly, we take a M -dimensional cubic region D of measuring range within a direct sum of the reachable sets as

$$\bigoplus_{i=1}^9 \Psi_i \supset D := [a_1, b_1] \times \cdots \times [a_M, \cdots, b_M] \quad (32)$$

where \bigoplus represents a direct sum (disjoint union) and $[a, b]$ denotes an interval. We divide it into a direct sum of uniform subcubes D^i as

$$D = \bigoplus \left\{ D^i \mid i \in \mathbb{I} = [1, \cdots, m_1] \times \cdots \times [1, \cdots, m_M] \right\}, \quad (33)$$

where \mathbb{I} is a space of M -dimensional indices and m_j is the number of subcubes in the j th direction. We then define center points of the subcubes $\mathbf{y}^i \in D^i$ ($i \in \mathbb{I}$) whose j th component is given by

$$(\mathbf{y}^i)_j := a_j + \left((i)_j - \frac{1}{2} \right) \frac{b_j - a_j}{m_j}, \quad (34)$$

where $(\mathbf{v})_j$ denotes the j th component of a vector \mathbf{v} .

The setup above allows us to build a numerical method as follows:

- (1) As an offline learning procedure, the mapping \bar{C} :

$$\bar{C}(i) := \nu \quad \text{if} \quad \lim_{t \rightarrow \infty} \phi_t(h(\mathbf{y}^i), T) = \omega_\nu, \quad (35a)$$

is numerically stored by solving (21) from $\xi_0 = h(\mathbf{y}^i)$.

- (2) Within the IC in process, the classifier C is quantized by C^* as

$$C(\mathbf{x}) \approx C^*(\mathbf{x}) := \bar{C}(i) \quad \text{for } i \text{ such that } H\mathbf{x} \in D^i. \quad (35b)$$

In this method, accuracy of the classifier C^* depends on the dimension of measurement M , the size and placement of measuring range $[a_j, b_j]$, and the resolution of quantization of reachable set m_j ($j = 1, \cdots, M$).

In summary, we have obtained the IC for the left-hand agent as a composed function of the quantized classifier C^* , the selector $S_{\mathbb{J}}$, and the impulse generator G , which results in the closed-loop form:

$$u_1^{\text{ic}} = u_1^{\text{ic}}(\mathbf{x}(t); \mathbb{J}) := P_1 \cdot (G \circ S_{\mathbb{J}} \circ C^*)(\mathbf{x}(t)). \quad (36)$$

If $M = 8$ and τ_G is sufficiently large, it is implied that the solution $\mathbf{x}(t)$ in (21) starting from any initial states in D undergoes an impulsive force at the time $t = t_0$ that u_1^{ic} decides autonomously and that it converges to the stable equilibriums specified by \mathbb{J} , under the resolution limit, $\min_j(m_j) \rightarrow \infty$.

3.6. Reduced-order design of measurement for rigid connection

The connection rod can behave like a rigid rod if the viscoelastic parameters k_w, c_w are sufficiently large. For the rest of this paper, we restrict our problem to this nearly rigid connection. Although human wrestling involves more flexible interactions between agents, this allows us substantially to reduce the computational efforts as follows. In this case, dependency between displacements and velocities of the left-hand and right-hand carts is imposed in the following form:

$$x_2 = x_2(x_1, \theta_1, \theta_2) = x_1 - r(\sin \theta_2 - \sin \theta_1) + \sqrt{w_0^2 - r^2(\cos \theta_2 - \cos \theta_1)^2}, \quad (37a)$$

$$\dot{x}_2 = \dot{x}_2(\dot{x}_1, \theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2) = \dot{x}_1 - r(\dot{\theta}_2 \cos \theta_2 - \dot{\theta}_1 \cos \theta_1) + \frac{r^2(\cos \theta_2 - \cos \theta_1)(\dot{\theta}_2 \sin \theta_2 - \dot{\theta}_1 \sin \theta_1)}{\sqrt{w_0^2 - r^2(\cos \theta_2 - \cos \theta_1)^2}}, \quad (37b)$$

Table 1 Parameter setting of the CIP model.

| Parameters | Values |
|----------------|---|
| m_θ | mass of pendulum 0.68 kg |
| m_x | mass of cart 0.067 kg |
| g | acceleration of gravity 9.8 m/s ² |
| r | length of pendulum 0.3 m |
| c_x | viscous coefficient along x 0.01 Ns/m |
| c_θ | viscous coefficient about θ 0.01 Ns/m |
| w_0 | natural length of connection rod 1 m |
| k_w | spring coefficient of connection rod 5000 N/m |
| c_w | viscous coefficient of connection rod 50 Ns/m |
| k_f | spring coefficient of floor 500 N/m |
| c_f | viscous coefficient of floor 10 Ns/m |
| μ | Coulomb friction coefficient of floor 0 |
| σ | steepness of step function 10 ⁶ |
| α | steepness of trapezoidal function 25 |
| K_p | proportional gain of standing control 1 |
| K_d | derivative gain of standing control 0.01 |
| $\Delta\theta$ | threshold of standing control $\pi/6$ rad |

Table 2 Parameter setting of the four-dimensional IC and numerical integration.

| Parameters | Values |
|-----------------|---|
| Q_{\max} | maximal strength of impulse 0.06 Nms |
| P_1 | strength of impulse of the 1st IC $= Q_{\max}$ |
| P_2 | strength of impulse of the 2nd IC $= -Q_{\max}$ |
| $\Delta\tau$ | width of impulse $= \Delta t$ |
| τ_G | relaxation time of impulse generator $= \Delta t$ |
| D | region of measurement $[-0.13, 0.43] \times [-3.28, 10.58] \times [-0.35, 0.31] \times [-3.80, 5.15]$ |
| Δt | step size of numerical integration 5×10^{-4} s |
| Free parameters | |
| m | resolution of numerical classifier ($m_j := m$ for all j) |
| Q | strength of impulsive disturbance |
| τ^d | delay time of delay element |

where $w = w_0 \gg 2r$ is a constant length of the rigid rod. Then, a loss-less state feedback to IC can be done by the following measurement:

$$\mathbf{y} = H\mathbf{x} = (x_1, \dot{x}_1, \theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)^T, \quad H = [\mathbf{e}_1^{(6)}, \mathbf{e}_2^{(6)}, \mathbf{e}_3^{(6)}, \mathbf{e}_4^{(6)}, \mathbf{o}^{(6)}, \mathbf{o}^{(6)}, \mathbf{e}_5^{(6)}, \mathbf{e}_6^{(6)}], \quad (38a)$$

where $\mathbf{e}_i^{(d)}$, $\mathbf{o}^{(d)}$ denote the i th standard basis vector and the zero vector in Euclidean space R^d respectively. A corresponding inverse satisfying the rigid constraint can be defined by

$$h(\mathbf{y}) := H^+ \mathbf{y} + x_2(x_1, \theta_1, \theta_2)\mathbf{e}_5^{(8)} + \dot{x}_2(\dot{x}_1, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_2)\mathbf{e}_6^{(8)}, \quad (38b)$$

where H^+ is the Moore-Penrose pseudoinverse of H . This measurement reduces computational efforts to obtain the quantized reachable sets of \bar{C} in (35a) from $O(m^8)$ to $O(m^6)$ with respect to the resolution of quantization m .

In the following numerical examples, we perform a further reduction of order given by

$$\mathbf{y} = H\mathbf{x} = (\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)^T, \quad H = [\mathbf{o}^{(4)}, \mathbf{o}^{(4)}, \mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{o}^{(4)}, \mathbf{o}^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}], \quad (39a)$$

$$h(\mathbf{y}) := H^+ \mathbf{y} + x_1\mathbf{e}_1^{(8)} + \dot{x}_1\mathbf{e}_2^{(8)} + x_2(x_1, \theta_1, \theta_2)\mathbf{e}_5^{(8)} + \dot{x}_2(\dot{x}_1, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_2)\mathbf{e}_6^{(8)} \quad \text{with } x_1 = \dot{x}_1 = 0. \quad (39b)$$

Although this measurement suffers complete loss of information about the cart motion x_1, \dot{x}_1 (and x_2, \dot{x}_2 via (37)), it reduces the computational efforts into $O(m^4)$. In this paper, we employ this four-dimensional measurement in priority to reducing the computational efforts. Moreover, in this four-dimensional measurement, the controller $u_1^{\text{ic}}(\mathbf{x}; \mathbb{J})$ originally designed for the left-hand agent can symmetrically be reused for the right-hand agent by a transformation:

$$u_2^{\text{ic}}(\mathbf{x}; \mathbb{J}') := u_1^{\text{ic}}(\mathbf{x}'; \mathbb{J}), \quad P_2 := -P_1, \quad \mathbf{x}' := -(\mathbf{x}_2^T, \mathbf{x}_1^T)^T, \quad (40)$$

where $\mathbb{J}' = \{4, 7\}$ for $\mathbb{J} = \{2, 3\}$ due to the transpose of the 3×3 matrix of ω_i in Fig. 3.

4. Numerical investigation

We conduct numerical experiments to evaluate performance of the four-dimensional IC of the measurement in (39). For simplicity, in what follows, we set the resolution m_j of quatization to a common value m for all j . The parameter values used in the following examples are listed in Table 1. The physical dimensions m_θ, m_x and r are roughly collected from the commercially available inverted pendulum (ZMP INC., 2011). For numerical integration, the fourth-order Runge-Kutta-Gill method is employed with the time step Δt listed in Table 2.

4.1. Individual performance of IC

In order to investigate individual performance of IC, we first consider impulse responses of the CIP model in (15) equipped with the IC in the left-hand only given by

$$\mathbf{T} = \mathbf{u}^{\text{pd}} + \mathbf{u}^{\text{ic}} + \mathbf{v} = \mathbf{u}^{\text{pd}} + \begin{bmatrix} u_1^{\text{ic}}(\mathbf{x}(t); \mathbb{J}_1) \\ 0 \end{bmatrix} + \begin{bmatrix} v(t) \\ 0 \end{bmatrix}, \quad v(t) := QI_{\Delta\tau}(t), \quad \mathbb{J}_1 := \{2, 3\}, \quad (41)$$

where \mathbf{u}^{pd} is the standing controller in (12), Q is an impulse of initial disturbance, \mathbf{u}_1^{ic} is the four-dimensional IC developed in (39), and $I_{\Delta\tau}(t)$ is the unit impulse function defined in (19).

In the following numerical examples, we assume the maximal strength of impulse $Q_{\max} = 0.06$ so that $v(t)$ cannot produce switching motions from the trivial initial state $\mathbf{x}(0) = \mathbf{o}^{(8)} + w_0 \mathbf{e}_5^{(8)}$ (or ω_1) to the other stable states in order to avoid trivial switching motions. Following this assumption, the region of measurement D is taken as listed in Table 2 so that it circumscribes at least all the trajectories $\mathbf{y}(t) = H\mathbf{x}(t)$ for the maximal disturbance $Q = Q_{\max}$. We also set the strength of impulse of IC to a common value $P_1 = -P_2 = Q_{\max}$ to avoid the trivial switching motions mentioned above. Note that under the conditions of k_w and c_w listed in Table 1 and for $Q \leq Q_{\max}$, dynamic change of length of the connection rod in (41) is limited to $|w(t) - w_0|/w_0 < 0.01$ so that the measurement in Section 3.6 is expected to work.

Solving (15) with (41) numerically from a trivial initial state $\mathbf{x}(0) = \mathbf{o}^{(8)} + w_0 \mathbf{e}_5^{(8)}$ for a given Q , we have the corresponding final position ω_v and obtain correspondence from Q to v as plotted in Fig. 7 for the resolution $m = 50$ where Q is taken at $N_Q = 100$ points with uniform increment over the interval $[0, Q_{\max}]$. The small circles represent the results of $v(Q)$ in presence of the IC's outputs and the cross marks represent those in absence of the outputs due to weak $v(t)$.

To evaluate the performance, we define a success rate in the following form:

$$E = E(\mathbb{J}) := N_{\mathbb{J}}/(N_Q - N_0) \quad (0 \leq E \leq 1), \quad (42)$$

where $N_{\mathbb{J}}$ is the number of points on Q at which $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \omega_v$ for all $v \in \mathbb{J}$ and N_0 is the number of points in absence of the IC's outputs. The success rate of the result for $m = 50$ shown in Fig. 7 is calculated as $E = 0.165$. Also, the rate for $m = 100$ can be obtained as $E = 0.305$ in the same manner. Therefore, it appears that our IC has low performance.

4.2. Performance improvement with a delay element

One of the reasons of the low performance above can be explained in Fig. 8. The colored areas represent the quantized reachable sets:

$$\Psi_v^* := \bigoplus \left\{ D^i \mid \bar{C}(i) = v, i \in \mathbb{I} \right\} \quad (v = 1, \dots, 9), \quad (43)$$

plotted on the hyper plane that contains the point $\mathbf{x}(t_1)$ on the phase trajectory at which the IC's output occurs (indicated by the cross mark). The resolutions of quantization of reachable sets are $m = 50$ for Fig. 8 (a) and $m = 1000$ for Fig. 8 (b). In the low resolution case in Fig. 8 (a), it can numerically be clarified that $\mathbf{x}(t_1)$ belongs to the reachable set Ψ_2^* (in red) of ω_2 while $\mathbf{x}(t)$ ($t > t_1$) actually converges to ω_4 ($\neq \omega_2$). Such misclassification can be refined by the high resolution as shown in Fig. 8 (b). In this case, the point $\mathbf{x}(t_1)$ is primarily classified into Ψ_4^* (in blue) of ω_4 . Although in theory, taking a sufficiently large resolution m provides nearly exact reachable sets, it greatly enlarges computational efforts. Another approach to reduce the misclassification is to replace the quantized reachable set Ψ_i^* in (43) of ω_i with a subset $\Psi_i^\circ := \Psi_i^* - \Delta\Psi_i^*$ where $\Delta\Psi_i^* \subset \Psi_i^*$ is a set of border points neighboring the other sets Ψ_j^* ($j \neq i$). However, extracting the border points is not necessarily easy because the reachable sets sometimes exhibit nested structures as discussed by one of the authors (Yoshida, 2009). Actually, in Fig. 8 (b), quite narrow region of Ψ_1^* (in white) appears between Ψ_2^* (in red) and Ψ_4^* (in blue).

Therefore, let us take yet another approach by replacing the IC with a delayed IC (DIC) in the following form:

$$\mathbf{u}_i^{\text{dic}}(\mathbf{x}(t); \mathbb{J}, \tau^d) := \mathbf{u}_i^{\text{ic}}(\mathbf{x}(t - \tau^d); \mathbb{J}) \quad (i = 1, 2), \quad (44)$$

where τ^d is a delay time. It seems reasonable to expect that a trajectory about to crossing a course-grained border of a reachable set reaches the true border soon. Figure 9 shows the success rate E corresponding to this replacement given by

$$\mathbf{T} = \mathbf{u}^{\text{pd}} + \begin{bmatrix} \mathbf{u}_1^{\text{dic}}(\mathbf{x}(t); \mathbb{J}_1, \tau^d) \\ 0 \end{bmatrix} + \mathbf{v}, \quad \mathbb{J}_1 := \{2, 3\}, \quad (45)$$

where E is averaged over the two types of initial disturbance: $\mathbf{v}(t) = (v(t), 0)^T$ and $(0, v(t))^T$, and the other procedures of obtaining E are the same as in Fig. 7. In Fig. 9, the triangles and the circles represent E as functions of the delay time τ^d of DIC in (44) for $m = 50$ and 100 respectively. It is clearly seen that the functions $E(\tau^d)$ are nearly concave down. The maximal values are $E = 0.392$ at $\tau^d = 0.0045$ and $E = 0.683$ at $\tau^d = 0.0025$ for $m = 50$ and 100 respectively. Therefore, the DIC roughly doubles the success rate, namely, $0.392/0.165 \approx 2.38$ for $m = 50$ and $0.683/0.305 \approx 2.24$ for $m = 100$.

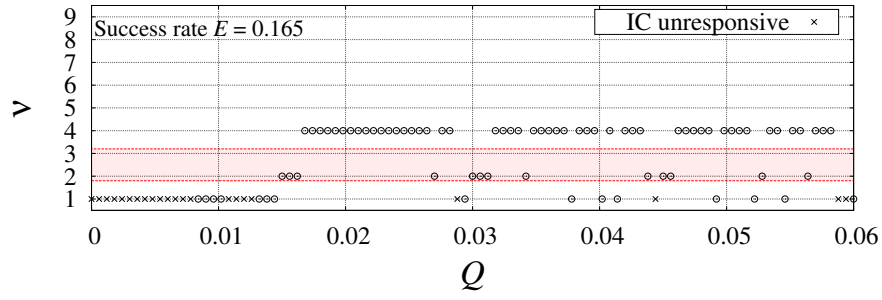


Fig. 7 Index of final state ω_v as a function of the strength of initial disturbance Q for the resolution $m = 50$.

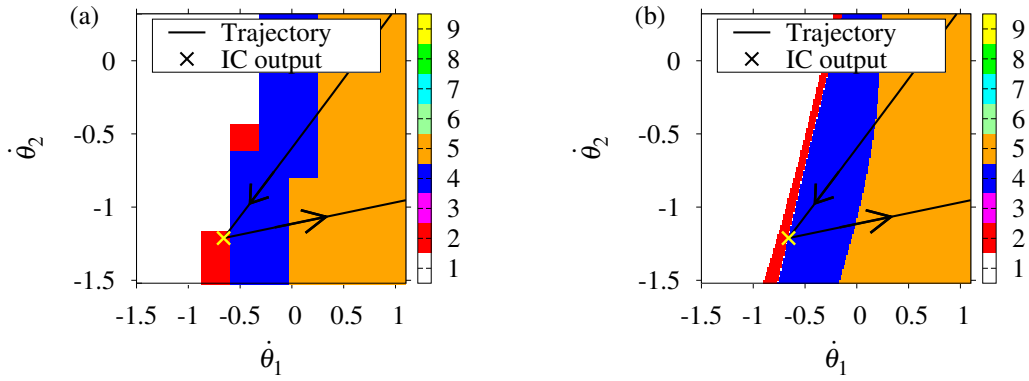


Fig. 8 Misclassification of reachable sets. Colored areas are quantized reachable sets for (a) $m=50$ and (b) $m=1000$.

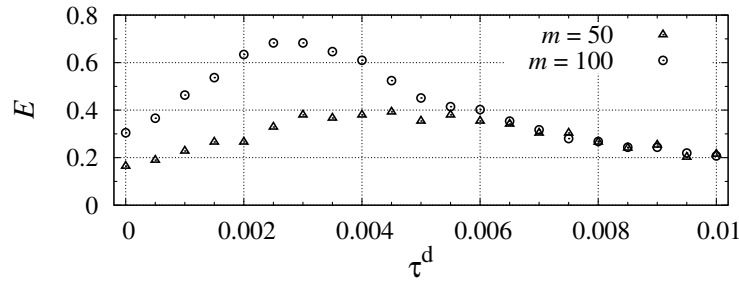


Fig. 9 Success rate E as functions of the delay time τ^d of DIC in (44).

4.3. Competitive behavior

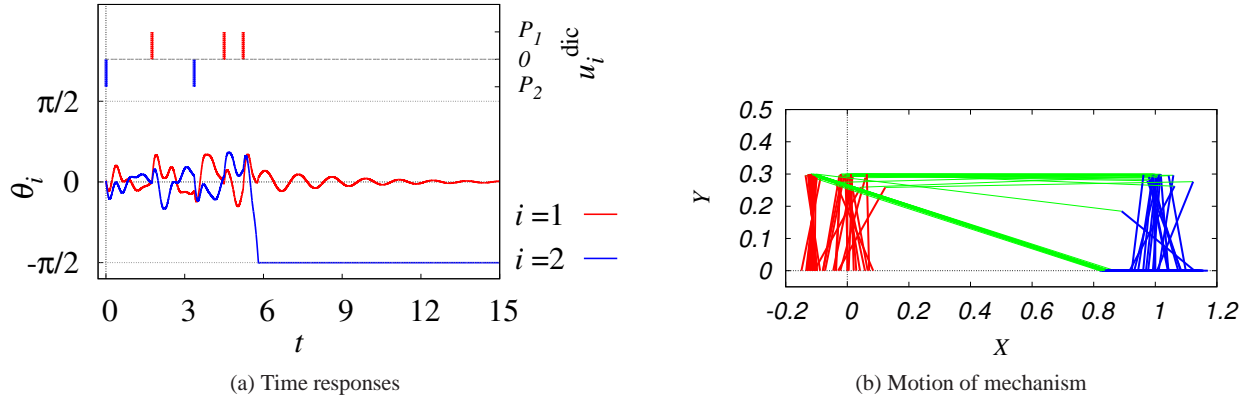
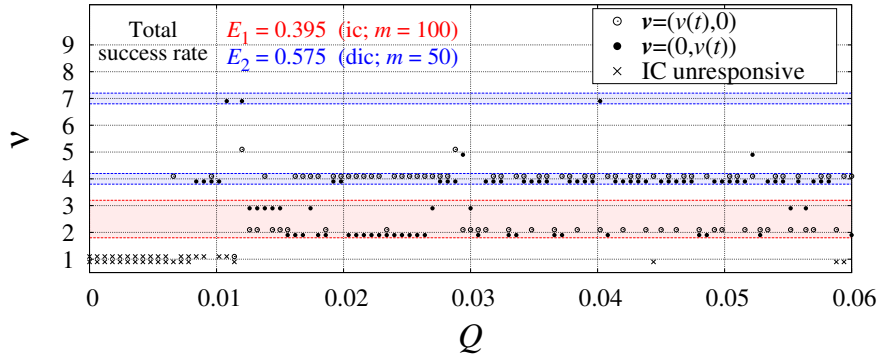
In this final section, we install the DICs into both sides of controllers as

$$\mathbf{T} = \mathbf{u}^{\text{pd}} + \begin{bmatrix} u_1^{\text{dic}}(\mathbf{x}(t); \mathbb{J}_1, \tau_1^d) \\ u_2^{\text{dic}}(\mathbf{x}(t); \mathbb{J}_2, \tau_2^d) \end{bmatrix} + \mathbf{v}, \quad \mathbb{J}_1 := \{2, 3\}, \quad \mathbb{J}_2 := \{4, 7\}, \quad (46)$$

where \mathbf{u}^{pd} is the standing controller in (12), u_1^{dic} and u_2^{dic} is the four-dimensional DIC in (44) through (40), and $\mathbf{v}(t) = (v(t), 0)^T, (0, v(t))^T$ are initial impulsive disturbances of strength Q as shown in (41). The competitive meanings of \mathbb{J}_1 and \mathbb{J}_2 are shown in Fig. 3.

Figure 10 shows a competitive behavior between the normal IC: $u_1^{\text{dic}}(\mathbf{x}(t); \mathbb{J}_1, 0) = u_1^{\text{ic}}(\mathbf{x}(t); \mathbb{J}_1)$ for $m = 100$ and the DIC with optimal delay time: $u_2^{\text{dic}}(\mathbf{x}(t); \mathbb{J}_2, 0.0045)$ for $m = 50$. The time responses in Fig. 10 (a) are obtained by solving (15) with (46) from $\mathbf{x}(0) = \mathbf{o}^{(8)} + w_0 \mathbf{e}_5^{(8)}$ under the conditions listed in Table 1 and Table 2 by applying the initial disturbance $\mathbf{v}(t) = (v(t), 0)^T$ with $Q = 0.0186$. The same solution is plotted in Fig. 10 (b) as a motion of the CIP mechanism. For convenience, we refer to the former IC as “IC100” and the later as “DIC50”.

It is shown in Fig. 10 that the ICs above generate the impulsive control forces autonomously to make the CIP system drop into ω_2 . According to the competitive interpretation in Fig. 3 inspired by some kind of wrestling match, it can be

Fig. 10 Competitive behavior for IC100 vs DIC50 ($\tau_2^d = 0.0045$).Fig. 11 Index of final state ω_v as functions of Q for IC100 vs DIC50 ($\tau_2^d = 0.0045$).

said that “the left agent wins by pulling the right agent.”

Figure 11 shows the result of competition. The plots are obtained in the same manner as those in Fig. 7, except that in Fig. 11, the open and filled circles represent the results for $v(t) = (v(t), 0)^T$ and $(0, v(t))^T$ respectively, and that the cross marks represent the results in which neither of IC100 and DIC50 produces its own output. Similar to the individual case in (42), we define the success rate of this competition as

$$E_i = E(\mathbb{J}_i) := N_{\mathbb{J}_i} / (N_Q - N_0) \quad (0 \leq E_i \leq 1) \quad (i = 1, 2), \quad (47)$$

where $N_Q = 100 \times 2$ is the number of all plots in Fig. 11, N_0 is the number of trials in absence of the IC's outputs (cross marks), and the definition of $N_{\mathbb{J}_i}$ is the same as that of $N_{\mathbb{J}}$ in (42). From the result in Fig. 11, the success rates are obtained as $E_1 = 66/167 \approx 0.395$ for IC100 and $E_2 = 96/167 \approx 0.575$ for DIC50. Therefore, it is shown that at least based on the present definition of competition and success rate, the performance improvement using time delay is more effective than that doubling the quantization resolution without the time delay.

5. Conclusion

In this paper, we discussed a competitive problem in which mechanical agents are fighting with each other and formulated it as the set of: (A) the nonlinear dynamical model with nine stable equilibriums and (B) the matrix describing competitive interpretation of these equilibriums. Based on this framework, we proposed a competitive IC that receives the state vector and output the impulsive forces to make the competitor fall down. Developing a quantized and reduced order design of the controller, we derived a practical control procedure along with an off-line learning method. To investigate performance of the controller in individual use and also in competition use, we conducted numerical experiments and obtained the following results:

- The individual performance of IC depends on resolutions of the quantized reachable sets.
- The individual performance of IC can be improved by adding the delay element into the IC.

- To improve the competitive performance of IC, adding the delay element may become more effective than refining the resolutions.

In future work, we plan to investigate a further order reduction of measurement based on time delayed embedding methods and to improve classification accuracy by applying machine learning techniques. We also plan to conduct competitions between humans and the proposed ICs. Moreover, application of position control to the carts will be considered to investigate competitive problems in a bounded area.

Acknowledgment

The authors appreciate the feedback offered by Dr. Munehisa Sekikawa.

References

- Bayadi, R., Banavar, R. N. and Chang, D. E., Characterizing the reachable set for a spacecraft with two rotors, *Systems & Control Letters*, Vol.62, No.6 (2013), pp.453–460.
- Duncan, P. W., Weiner, D. K., Chandler, J. and Studenski, S., Functional reach: A new clinical measure of balance, *Journal of Gerontology*, Vol.45, No.6 (1990), pp.M192–M197.
- Hofbauer, J. and Sigmund, K., *Evolutionary Games and Population Dynamics* (1998), Cambridge University Press.
- Hsu, P., Coordinated control of multiple manipulator systems, *IEEE Transactions on Robotics and Automation*, Vol.9, No.4 (1993), pp.400–410.
- Lhommeau, M., Jaulin, L. and Hardouin, L., Capture basin approximation using interval analysis, *International Journal of Adaptive Control and Signal Processing*, Vol.25, No.3 (2011), pp.264–272.
- Maravall, D., de Lope, J. and Domínguez, R., Coordination of communication in robot teams by reinforcement learning, *Robotics and Autonomous Systems*, Vol.61, No.7 (2013), pp.661–666.
- Moore, M. and Wilhelms, J., Collision detection and response for computer animation, *SIGGRAPH Computer Graphics*, Vol.22, No.4 (1988), pp.289–298.
- Nakamura, Y., Nagai, K. and Yoshikawa, T., Mechanics of coordinative manipulation by multiple robotic mechanisms, *Proceedings of 1987 IEEE International Conference on Robotics and Automation*, Vol.4 (1987), pp.991–998.
- Nelson, A., Grant, E. and Henderson, T., Evolution of neural controllers for competitive game playing with teams of mobile robots, *Robotics and Autonomous Systems*, Vol.46, No.3 (2004), pp.135–150.
- Pagello, E., D'Angelo, A., Montesello, F., Garelli, F. and Ferrari, C., Cooperative behaviors in multi-robot systems through implicit communication, *Robotics and Autonomous Systems*, Vol.29, No.1 (1999), pp.65–77.
- Panwar, V., Kumar, N., Sukavanam, N. and Borm, J., Adaptive neural controller for cooperative multiple robot manipulator system manipulating a single rigid object, *Applied Soft Computing*, Vol.12, No.1 (2012), pp.216–227.
- Stone, P. and Veloso, M., Multiagent systems: A survey from a machine learning perspective, *Autonomous Robots*, Vol.8, No.3 (2000), pp.345–383.
- Wu, J., Snášel, V., Ochodková, E., Martinovič, J., Svatoň, V. and Abraham, A., Analysis of strategy in robot soccer game, *Neurocomputing*, Vol.109 (2013), pp.66–75.
- Yoshida, K., Fractal dependence on initial conditions of coupled inverted pendula model of competition and cooperation, *Journal of System Design and Dynamics*, Vol.3, No.6 (2009), pp.966–974.
- Yoshida, K. and Ohta, H., Coupled inverted pendula model of competition and cooperation, *Journal of System Design and Dynamics*, Vol.2, No.3 (2008), pp.727–737.
- ZMP INC., e-nuvo WHEEL (2011), <http://www.zmp.co.jp/e-nuvo/en/>.